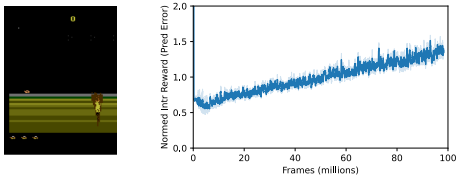
Neuro-Inspired Fragmentation and Recall to Overcome Catastrophic Forgetting in Curiosity

Jaedong Hwang Zhang-Wei Hong Eric R Chen Akhilan Boopathy Pulkit Agrawal Ila Fiete

Massachusetts Institute of Technology

❑ Catastrophic Forgetting is not only issue in Continual Learning

▪ Intrinsic Reward of Same Observation

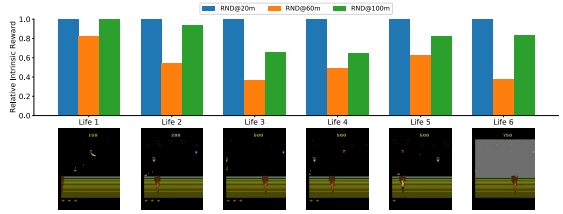


▪ Key Assumption of Prediction-Based Intrinsic Reward Function

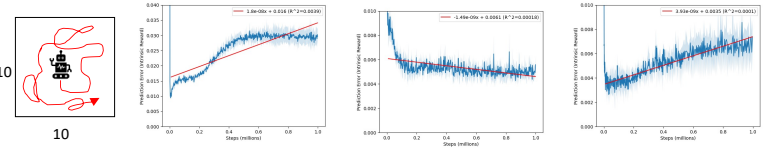
$$r^i_{t_1}(s) \leq r^i_{t_2}(s) \iff t_1 > t_2$$

Monotonically decreasing function

▪ Intrinsic Reward before Dying during Training



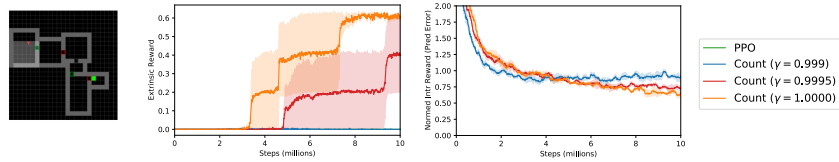
▪ Distribution Shift Occurs as the Agent Explore More Space



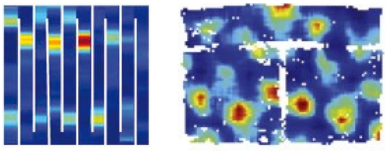
▪ Catastrophic Forgetting Deteriorates Performance

$$N_{\text{visit}}(o) \leftarrow \gamma \cdot N_{\text{visit}}(o) + b$$

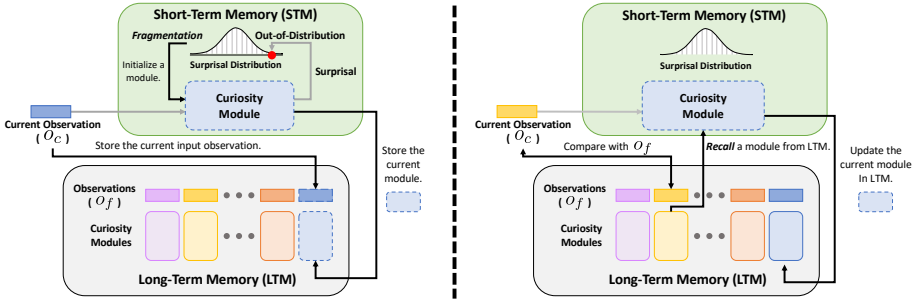
$$r_i = 1/\sqrt{N_{\text{visit}}(o)}$$



❑ Fragmentation and Recall in Curiosity-Driven RL (FARCuriosity)

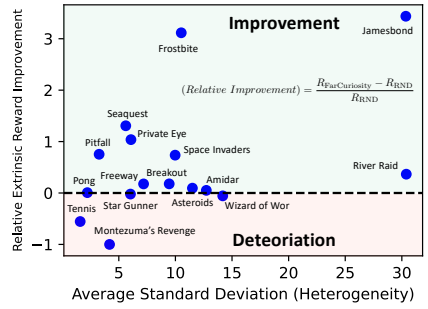


Derdikman et al. (2009) Carpenter et al. (2015)



- **Grid cell** in entorhinal cortex divides an environment into multiple fragments.
- **STM** holds current curiosity module (e.g., **RND**)
- **LTM** stores curiosity modules and observations on fracture points.

▪ Comparison with RND (Burda et al., 2019)



- **Relative Improvement**

$$\frac{R_{\text{FarCuriosity}} - R_{\text{RND}}}{R_{\text{RND}}}$$

R_* : extrinsic reward

- **Heterogeneity**
  Average standard deviation of each pixel in generated videos from trained RND.

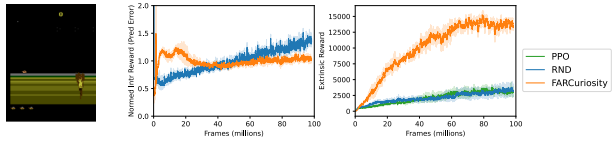
▪ Fragmentation

- The prediction error in curiosity module is used for the surprisal.
- Surprisal > Out-of-Distribution of running statistics.
- Store the current observation and the curiosity module in LTM.
- Initialize a new curiosity module.

▪ Recall

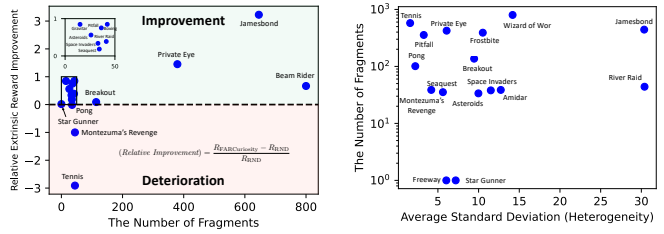
- Cosine similarity between the current observation and fracture point's observation is high, store the current module and recall the corresponding one.

▪ Normalized Intrinsic Reward of First Frame in Jamesbond



- Intrinsic reward from **RND** keeps increasing ⇒ **Catastrophic Forgetting**
- **FarCuriosity** generates consistent reward after 20M steps.

▪ No Global Trend between N(fragments) and the Performance



❑ **Reference** Burda, et al., Exploration by random network distillation. *ICLR 2019.* | Derdikman et al. Fragmentation of grid cell maps in a multicompartment environment. *Nature Neuroscience. 2009.* | Carpenter, et al. Grid cells form a global representation of connected environments. *Current Biology, 2015.*